VOICE-BASED SCREEN NAVIGATION APPARATUS AND METHOD

CROSS-REFERENCE TO RELATED APPLICATION(S)

[0001] This application claims the benefit under 35 USC 119(a) of Korean Patent Application No. 10-2015-0107523, filed on Jul. 29, 2015, in the Korean Intellectual Property Office, the entire disclosure of which is incorporated herein by reference for all purposes.

BACKGROUND

[0002] 1. Field

[0003] The following description relates to a voice-based screen navigation apparatus and method.

[0004] 2. Description of Related Art

[0005] There have been suggested apparatuses of various forms and methods for checking information that is displayed on devices, such as TVs, computers, tablet devices, and smartphones, as well as separately for inputting commands to process the checked information. Generally, information is input via a device, such as a remote controller, a mouse, and a keyboard, or via touch input. More recent input attempts have involved the interpreting of user voice input to control a device. However, these latest attempts only enable execution of implementing designated functions or simple applications based on preset, fixed commands.

SUMMARY

[0006] This Summary is provided to introduce a selection of concepts in a simplified form that are further described below in the Detailed Description. This Summary is not intended to identify key features or essential features of the claimed subject matter, nor is it intended to be used as an aid in determining the scope of the claimed subject matter.

[0007] In one general aspect, a screen navigation apparatus includes a command receiver configured to receive an input voice command regarding navigation of a screen, and a processor configured to interpret the voice command based on an analysis result of content displayed on a screen and compose a command executable by the screen navigation apparatus and perform navigation of the screen.

[0008] The screen navigation apparatus may further include a memory configured to store instructions. The processor may be further configured to execute the instructions to configure the processor to interpret the voice command based on the analysis result of the content displayed on the screen and compose the command executable by the screen navigation apparatus, and perform the navigation of the screen.

[0009] The processor may include a command composer configured to interpret the voice command based on the analysis result of the content displayed on the screen and compose the command executable by the screen navigation apparatus, and a command executer configured to perform the navigation of the screen.

[0010] The processor may further include a screen analyzer configured to analyze the content displayed on the screen and generate the content analysis result. The screen analyzer may be configured to analyze the content using one or more of the following techniques: source analysis, text analysis, speech recognition, image analysis and context information analysis. The content analysis result may include a semantic map or a screen index, or both, wherein the semantic map represents a determined meaning of the content displayed on the screen, and the screen index indicates a determined position of the content displayed on the screen. The screen index may include at least one of the following items: coordinates, grids, and identification symbols, and the screen analyzer determines at least one of a type, size, and position of the screen index to be displayed on the screen by taking into account at least one of the following factors: coordinates of the screen index, a screen resolution, and positions and distribution of key contents on the screen, and displays the screen index on the screen based on the determination. In response to a user selecting one of screen indices displayed on the screen by a user's speech, eye-gaze, or gesture, or any combination thereof, the command composer may be configured to interpret the voice command based on screen position information that corresponds to the selected screen index.

[0011] The command receiver may be configured to receive the input voice command from a user in a predetermined form or in a form of natural language. The processor may further include the command receiver. The command composer may include a command converter configured to refer to a command set database (DB) and convert the input voice command into a command executable by the screen navigation apparatus. The command set DB may include a common command set DB or a user command set DB, or both, wherein the common command set DB stores common command sets and the user command set DB that stores command sets personalized for a user.

[0012] The command composer may include an additional information determiner configured to determine whether the input voice command is sufficient to be composed into the command, and a dialog agent configured to present a query to request the user to provide additional information in response to the determination indicating that the voice command is not sufficient. The dialog agent may be configured to create the query as multistage subqueries, and present a subquery based on a user's reply to a subquery presented in a previous stage.

[0013] The command composer may be configured to interpret the incoming voice command in stages and compose a command for each stage while the user's voice command is being input, and the command executer may be configured to navigate the screen in stages by executing the commands.

[0014] The navigation of the screen may include one or more of the following operations: keyword highlighting, zoom-in, opening a link, running an image, playing video, and playing audio.

[0015] The screen navigation apparatus may be a smartphone, a laptop, a tablet, a smart watch, or a computer, and may further include a screen and a user interface.

[0016] In another general aspect, a screen navigation method includes receiving a voice command regarding navigation of a screen, interpreting the voice command based on an analysis result of content displayed on the screen and composing a command, and performing navigation of the screen based on execution of the command.

[0017] The screen navigation method may further include analyzing content displayed on the screen and generating a content analysis result. The content analysis result may include a semantic map or a screen index, or both. The semantic map may represent a determined meaning of the